Version 2.2
November 10, 2000

User Facility Sub-Project Hardware and Material Costs

SUMMARY

This is a costing estimate of the User Facility subproject of the CMS S&C project. The first part of the document describes the scope of the User Facility project. Pricing of the various parts comes mainly from CDF/D0 Run 2 experience at Fermilab. The price evolution comes from Moore's laws, slopes as measured at CERN and fixed at endpoints from Fermilab measurements. Costs for R&D, networking and CMS Detector Construction phase support are also included.

THE CMS DATA ANALYSIS MODEL

The LHC distributed data analysis model has been described in [1]. Here we will summarize from this document some important terms and describe the size estimate for the Tier 1 and Tier 2 Regional Centers.

Definition of Terms

Raw Event

This refers to a raw data event as written by the online system. The size of the raw data event is about 1 MB.

Raw Simulated Event

This refers to a data event as simulated by the Monte Carlo and reconstruction packages. The size of the raw simulated event is about 2 MB .

AOD (Analysis Object Data)

refers to objects which facilitate analysis, but which by construction are not larger than 10 Kbytes. So, if an analysis requires information that does not fit in this limited size, this analysis should access other objects (possibly larger in size). A more traditional term for a collection of such objects is "micro-dst" or "ntuple".

Calibration Data

refers to diverse kinds of information generally acquired online (which may be strictly calibration constants, or monitoring data) as well as special runs taken for calibration purposes. Cosmic ray runs are included in this category. Collections of normal event data used for alignment and calibration studies are not classified as calibration data.

Data Caching

refers to the capability of holding a copy of frequently accessed data on rapidly accessible storage medium, under algorithmic control to minimize data turnaround and maximize user response.

ESD (Event Summary Data)

refers to physics objects of about 400 Kbytes. A more traditional term for this kind of object would be "mini-dst".

Tags

refer to very small objects (1 kB) that identify (tag) an event by its physics signature. A Tag could, for example, be a set of 96 bits each tagging a given physics channel or it could be a set of 10 words with some packed overall information for an event. In this document the Tag is NOT a global object identifier (i.e. a set of bits identifying uniquely an object in a database). A more traditional term would be "nano-dst" or "ntuple" The border between AOD and Tags is well defined in size but less well defined in functionality.

Data Analysis Tasks

The offline software of each experiment is required to perform the following tasks:

- data reconstruction, which may include more than a single step, such as preprocessing, reduction and streaming. Some of these steps might be done online;
- Monte Carlo production, which includes event generation, detector simulation and reconstruction;
- offline (re)calibration;
- successive data reconstruction; and
- physics analysis.

Basic Assumptions

We assume that there exists one "central site", CERN, which is able to provide all the various services (but not with enough total capacity to do all analysis related tasks).

The following steps happen at the central site only:

- Online data acquisition and storage
- Possible data pre-processing before first reconstruction
- First data reconstruction

Other production steps (calibration data storage, creation of ESD/AOD/tags) are shared between CERN and the Regional Centers.

The central site holds:

- a complete archive of all raw data
- a master copy of the calibration data (including geometry, gains, etc.)
- a complete copy of all ESD, AOD, and Tags preferably online

For a typical large LHC experiment, the data-taking estimate is:

- 1 PB raw data per year per experiment
- 10^9 events (1 MB each) per year per experiment
- 100 days of data taking (i.e. 10^7 events per day per experiment)

Size of the CMS T1 Regional Center at Fermilab

Total CPU needs at the Regional Centers
Here we break down by function the total CPU needs at the Tier 1 and Tier 2 centers in order to come up with a sizing estimate.

Simulation:
We estimate from physics requirements that US-CMS will need to simulate 10^8 events/year and each event takes about 5k Si95-seconds. We use the metric 2×10^7 seconds/year to give us an overall efficiency factor of 65% for CPU usage. This gives a total simulation CPU need of 25k Si95's. Most of this is done at Tier II sites, however some provision is made for smaller simulations at the Tier 1 Regional Center.

Reconstruction of simulated events:
We will need to reconstruct the 10^8 simulated events/year. Each event takes 3000 Si95-sec to reconstruct and using 65% efficiency this gives us 15k Si95 total.

Re-Reconstruction:
At the Tier 1 Regional Center we anticipate re-reconstructing 10% of the total raw data twice during the year, for a total of 10^8 events. Each event takes about 3k Si95-secs to reconstruct and using the 65% efficiency we get 30k SI95's total.

Event Selection:
We assume we will need to process the ESD samples to select and tag events. The ESD sample is about 10% of the full data sample and the processing will be done by 10 physics groups 6 times/year. The data selection job on the ESD takes 50 Si95-secs/event and using a 65% efficiency we get 15k Si95's needed for this task.

Physics Object Creation:
This is a scheduled task that reads ESD data and recreates the AOD physics objects with better calibrations and algorithms. Here we assume that 10 physics groups will want to do this each month and that each production covers 10^7 events. Using a 65% efficiency factor we get 12k Si95's needed for this task.

Derived Physics Data (ntuple) Creation:
These are scheduled jobs reading AOD samples and preparing tag/ntuples for detailed analysis by individual physicists. We estimate that 10 groups will read 10^7 events 20 times per year in this activity, and that each event requires 50 Si95-secs to process. Using the 65% efficiency gives us 5k Si95's required for this task.

Individual Analysis:
This task covers chaotic jobs from individual physicists reading AOD and DPD samples to prepare their own private DPD's for further analyses. Here we assume 200 physicists look at 10^7 events 20 times/year and that the analysis on each event takes 20 Si95. Because the jobs are unpredictable, we assume the CPU efficiency is half of the scheduled jobs (i.e. 33% efficient) providing capacity to handle the relatively common peaks normally seen in CPU usage. This gives us a requirement of 80k Si95s at Tier 1. We assume the Tier 2 centers will support about 20% of the physicists supported at Tier 1, so each Tier 2 center gets 16k Si95s.

The table below summarizes the amount of CPU needed at the Tier 1 and Tier 2 Regional Centers.

	Tier 1 Regional Center	**Tier 2 Regional Center (each)**
Task	**Total CPU**	**Total CPU**
Simulation	5k Si95	6k Si95
Reconstruction of Simulation	5k Si95	6k Si95
Re-reconstruction of raw data	30k Si95	
Event Selection	15k Si95	
Physics Object Creation	12k Si95	
Derived Physics Data Creation	5k Si95	
Individual Analysis	80k Si95	16k Si95
Data Servers (0.2×CPU)	15k Si95	
Total	**167k Si95**	**28k Si95**

Total Online Storage needs at the Regional Centers

Below we list the data we believe needs to be kept online. Because of performance issues or redundancy needs, we also assign an efficiency to disk depending on type of data stored on it. Most disk gets a 50% efficiency factor. The mass storage data cache disk is much more straightforward to manage and gets a 75% efficiency factor.

Simulated ESD/AOD
This refers to the ESD and AOD objects created by the simulation. Although most of the simulation happens at Tier 2 sites, the Tier 1 Regional Center will keep all of the ESD and AOD from the simulation on disk. 10^8 events X 400kB/event yields 20 TB for the simulated ESD. Having 2 versions of the AOD dataset gives us an additional 10 TB in AOD.

Calibration Data
The exact size and number of calibration events is not yet known, however we are reserving 10 TB in online disk space for calibration data. CMS is currently defining calibrations and the size of this data.

ESD data
We want to keep online at the Tier 1 center two versions of the ESD which means 400kB x 10^7 events x 2 = 80 TB.

AOD, TAG, etc.
These are the AOD and TAG data sets coming from physicist analysis. We imagine each AOD set to be on the order of 5TB and each TAG set being on the order of 0.5 TB.

Mass Storage Cache
This is disk used in front of the Mass Storage system to cache active data. It is estimated at about 20% of the 1 PB active data store.

AOD/Raw/ESD
This is a global estimate for online disk at a typical Tier 2 center.

User Disk
This is disk used by individual physicists or small workgroups to store their TAG or DPD event samples. For Tier 1 we estimate we need about 50 TB for this purpose and at Tier 2 sites we estimate about 25 TB.

	Tier 1 Regional Center		**Tier 2 Regional Center (each)**	
Task	**Total Disk**	**Efficiency**	**Total Disk**	**efficiency**
Simulated ESD/AOD	30 TB	50%		
Calibration Data	10 TB	50%		
2 ESD Versions	80 TB	50%		
AOD, TAG, etc.	20 TB	50%		
Mass Storage Cache	200 TB	75%		
AOD/raw/ESD			50 TB	50%
User Disk	50 TB	50%	25 TB	50%

Total	**390 TB**	**650 TB**	**75 TB**	**150 TB**

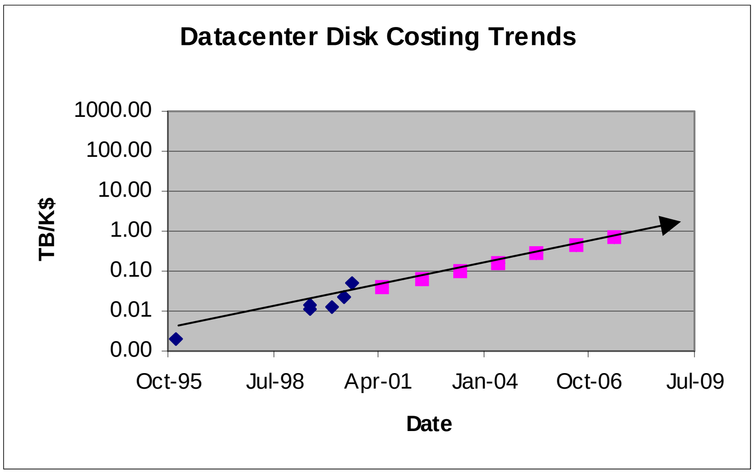
Total Offline Storage needs at the Regional Centers

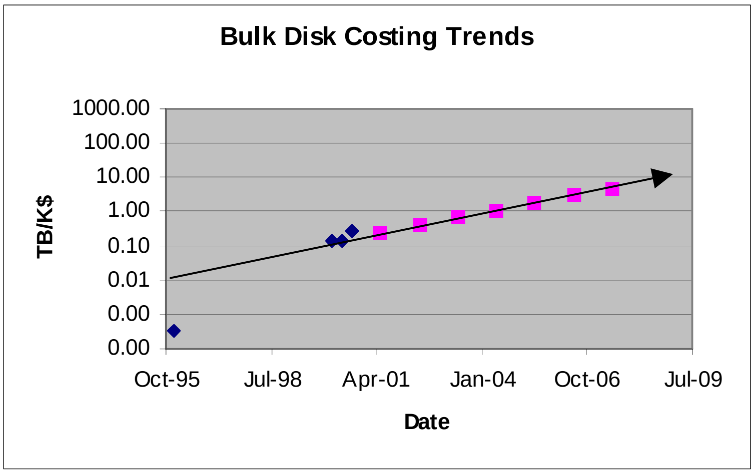
BASIS OF THE REGIONAL CENTER COST ESTIMATE

Assumptions of the Cost Estimates

Hardware costs for computing come from FNAL experience over the last few years in purchasing for Run 2 computing. We take the slopes of the extrapolations as measured at CERN [2] using Femilab data over the past few years as boundary conditions. Because much of the hardware is not going to be purchased for several years, and because of the dramatic drop in computing costs over time, an extrapolation is used. The assumptions for cost evolution are given below.

Disk Storage
There are two types of disk: bulk commodity disks and datacenter quality disks. The plots below show Run 2 disk purchases and the slope of the lines are the assumption that disk are assumed to halve in cost every 1.4 years. The pink points (>2001) are extrapolated.





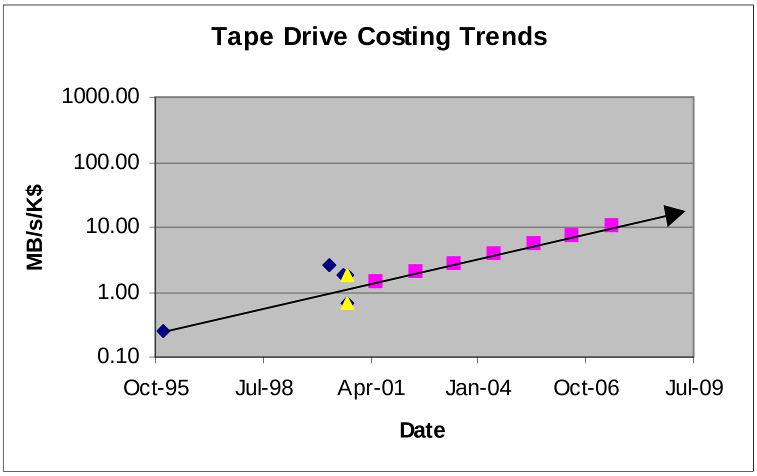
The extrapolated points in the above figures are listed in the table below.

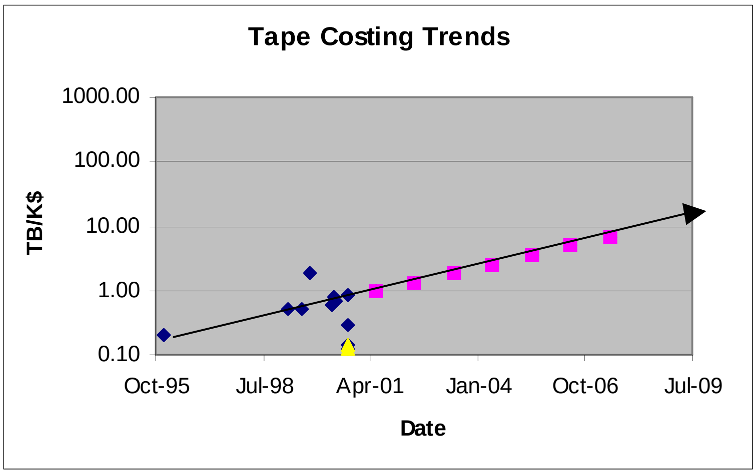
Year	2001	2002	2003	2004	2005	2006
Bulk Disk	$6.60/GB	$4/GB	$2.50/GB	$1.50/GB	$0.90/GB	$0.56
Datacenter Disk	$43.50GB	$26.50/GB	$16.20/GB	$9.90/GB	$6/GB	$3.70

Table 1: Extrapolated $/GB for bulk (commodity) disk and for data center quality disk.

Tape Storage
Here the assumptions are that robotics cost $75/slot, constant in time, and that tape density doubles every 2.1 years, implying that the robotics cost ½ every 2.1 years. The cost of tape media is assumed to ½ every 2.1 years. Tape drive costs/MB/sec are also expected to ½ every 2.1 years.

In the following plots, the pink (>2001) points are extrapolated, and the yellow points are for DVD media and drives. The blue points are traditional tape media and drives.





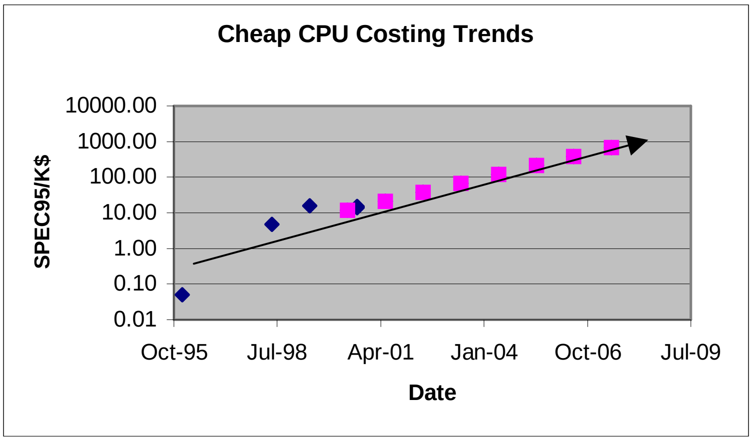
Year	2001	2002	2003	2004	2005	2006
Tape Drives	$943/MB/sec	$677/MB/sec	$487/MB/sec	$350/MB/sec	$252/MB/sec	$181/MB/sec
Media Costs	$1460/TB	$1050/TB	$760/TB	$540/TB	$390/TB	$280/TB

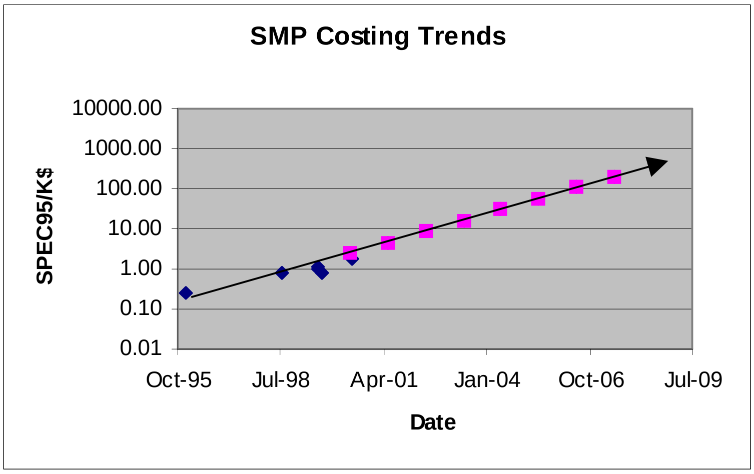
Table 2: Costs of tape media and drives.

CPU

Commodity CPU costs are estimated by assuming the performance of the CPU doubles every 1.2 years and the price stays the same. The starting number is from Run 2 farm purchases in June 1999, where 2060 Si95's were bought for $132,000 or 15.61 Si/$1000. The canonical dual CPU box was ~30 Si95's and cost $2000.

SMP CPU costs are estimated by assuming the performance of the CPU doubles every 1.1 years and the price stays the same. The starting point is for an SGI general use machine bought in 1998.





Year	2001	2002	2003	2004	2005	2006
Commodity CPU	$85/Si95	$48/Si95	$27/Si95	$15/Si95	$8.40/Si95	$5/Si95
SMP CPU	$410/ Si95	$220/Si95	$116/Si95	$61/Si95	$33/Si95	$18/Si95

Table 3: CPU costs.

HARDWARE COST BY WBS ITEM FOR THE USER FACILITY PROJECT

Here we detail by WBS number the hardware costs for the User Facility.

1.1 Tier 1 Regional Center

1.1.1 Development and Test systems

This item covers test systems needed to test both hardware and software used in the User Facility.

1.1.1.1 Modular Test Stands

These are hardware test stands typically on expert desktops for testing new hardware such as disks, tape drives, tapes and software drivers. The numbers here are a guess based on a test stand cost of typically $5000 and assuming we will need four of them each year.

2004	2005	2006	Total	2007 (ops)
$20,000	$20,000	$20,000	$60,000	$15,000

1.1.1.2 Integration Test Stands

This is the cost for a fully integral test stand to test production software and hardware. Most likely this will be the fully functional prototype regional center hardware with some upgrades. The cost here is some incremental costs to upgrade or retrofit the prototype regional center with any new technologies necessary.

2004	2005	2006	Total	2007 (ops)
$10,000	$10,000	$20,000	$40,000	$10,000

1.1.2 Simulation & Reconstruction systems

Assume the configuration from the Hoffman Review Panel 3. For costing we are staging by buying 10% in 2004, 30% in 2005 and 60% in 2006. Total here is installed capability in 2006.

Task	**CPU (Si95)**
Simulation	5,000
Reconstruction of Simulation	5,000
Re-reconstruction of raw data	30,000
Event Selection	15,000
Physics Object Creation	12,000
Derived Physics Data Creation	5,000
Total	**72,000**

The table below uses the extrapolated commodity computing costs, the total reconstruction/simulation CPU as detailed in the table above, and assumes a 10%, 30%, 60% staging over the years 2004-2006.

2004	2005	2006	Total	Operations
$108,055	$181,645	$203,889	493,589	164,000

1.1.2.1 Disk

There must be a disk pool to support the simulation work. This is not a large pool and is not used to store data. 1 TB should be fine. Implementation is all at the beginning and the disk is replaced every three years starting in 2007. These are data center disks. The efficiency of the disk is assumed to be 80% to cover RAID or need for increased performance.

2004	2005	2006	Total	Operations
$12,322			$12,322	$2,786

1.1.2.2 Hardware Infrastructure Costs

This is the infrastructure and other costs for the facility (racks, cooling, consoles, etc.) Using combined experience with Run 2 and more recent totally integrated purchases, we put this at about 10% of the total hardware costs. Because we used the more expensive data center disks for this, we didn't include the disk cost in this overhead.

2004	2005	2006	Total	Operations
$10,805	$18,164	$20,388	$49,357	$16,400

1.1.3 Analysis Systems

The table below uses commodity computing extrapolated prices to cost the 80,000 Si95's needed for user analysis as detailed above.

	2004	2005	2006	Total	2007 (operations)
CPU (Si95)	8,000	24,000	48,000	80,000	
Cost	$120,061	$201,827	$226,544	$548,432	$184,000

1.1.3.1 Hardware Infrastructure Costs

Because we think that building a cluster of commodity computing for chaotic physics analyses is unknown in how to serve data, how to load balance, and if special hardware will be necessary to integrate the cluster, we put a 50% integration cost on top of the hardware cost. This is the same assumption as made by experts at CERN.

2004	2005	2006	Total	Operations
$60,030	$100,913	$113,272	$274,216	$92,000

Note the networking costs are separated out in 1.6.1.

1.1.4 Data Access and Distribution Servers

This covers multiple AMS servers to serve the CMS database to/from the analysis & simulation systems. We use for a measurement the CPU usage for Objectivity which has been measured at ~ 20% overhead by CMS and BaBar. This overhead is calculated for all the CPU before any efficiency factors. Because the measurement included efficiencies, we don't add any additional efficiencies to this CPU, however we assume these servers are expensive SMPs. The operations costs assume a replacement every 3 years.

	2004	2005	2006	Total	2007 (operations)
Si95's	1500	4500	9000	15000	
cost	$929,541	$494,146	$263,143	$1,686,830	$562,280

1.1.5 Calibration and Non-Event Database Servers

Here we assume we need about 5% of the reconstruction Si95 and that this is also a high-end server class machine. The operations costs assume a replacement every 3 years.

	2004	2005	2006	Total	2007 (operations)
Si95's	360	1080	2160	3600	
cost	$223,089	$118,595	$63,154	$404,838	$134,946

1.1.6 Data Storage Systems

This includes disk and tape. The costs below are just for the disks and/or tapes, not for the server machines to serve the data (which is covered in 1.2.4).

Tape estimates

Task	**Tape (TB)**
Simulated event data	200
Genearl ESD' s	200
Simulated ESD+AOD for analysis	40
Raw Event Data	50
Reprocessed and Revised ESDs	500
AOD,TAG,Calibration,etc	30
Total	**1,020**

Tape is assumed to be 100% efficient.

Disk estimates

Task	Disk (TB)	Efficiency (%)	Total Disk
Simulated ESD+AOD for	30	50	60
Calibration Data	10	50	20
ESD for Local Analyses	80	50	160
AOD,TAG,etc	20	50	40
Mass Storage Cache	200	75	270
User Disk	50	50	100
Total	**390**		**650**

1.1.6.1 Data Storage Costs

In the below table, the online disk was priced by assuming an average between commodity and datacenter disk. The cost/GB for commodity disk was doubled to account for infrastructure costs.

Robotics cost was placed at $75/slot, which is the current cost for robotics. This price is not expected to go down (it's basically a cost for a mechanical system), however the number of slots needed goes down depending on tape density. We assume we will purchase a robot that can use commodity drives and tape media.

Robotic maintenance is currently $4.80/slot and is not expected to change.

	2004	2005	2006	Total	2007 (Operations)
Disk	417,822	762,961	930,060	2,110,843	703,614
Media	54,315	117,032	168,262	339,609	84,902
# Robotic Slots	3,250	3,250	0	6,500	
Robotics cost	243,750	243,750	0	487,500	97500
Robotics Maintenance	0	15,600	31,200	46,800	31,200
# Tape Drives	5	15	15	35	8
Tape Drives	50,000	150,000	150,000	350,000	80,000
Total	**765,887**	**1,289,343**	**1,279,522**	**3,334,752**	**997,216**

1.1.7 Data Import/Export

1.1.7.1 Tape import/export

We assume 100 tapes in/out per year. This is very uncertain. Will networking be good enough? Will we need more data import/export facility? This is a very modest sized system. We are assuming we can use the AML/J we have now throughout the lifetime of the experiment. Tapes are assumed to cost $75 each. Year 1 (2004) will start with 50 tapes/year and we go to 100 tapes/year after that. Tapes in year 2004 will hold on the order of 400 GB by our extrapolations which would mean exporting ~500 TB/year by tape.

Task	2004	2005	2006	Total	2007 (ops)
Tapes	$3,750	$7,500	$7,500	$18,750	$7,500
Maintenance of AML/J	$5,000	$5,000	$5,000	$15,000	$5,000
Refitting AML/J, drives, etc.	$15,000	0	$15,000	$30,000	$10,000
Total	$23,750	$12,500	$27,500	$63,750	$22,500

1.1.7.2 Tape import/export

There are no explicit hardware costs related to this item.

1.1.8 Data and Physics Analysis Software

There are no explicit hardware costs associated with this item.

1.1.9 System Administration

There are no explicit hardware costs associated with this item.

1.2 System and User Support

1.2.1 Documentation - document generation and access tools.

There are no explicit hardware costs associated with this item.

1.2.2 Collaborative Tools

1.2.2.1 Video conferencing

For video conferencing we need a couple of video conferencing rooms and some personal video conferencing equipment. We assume that personal video conferencing equipment (such as cameras on desktop pc's, etc) is something that comes with the encumbered salary. (i.e. VRVS is "free" with a desktop.)

In order to support the ~ 100 FNAL resident users in 2006 and the myriad of visitors, we estimate we will need 3 video conference rooms

	2001	2002	2003	2004	2005	2006	Total	2007 (operations)
Video rooms		1 room	1 room		1		3	upgrades
costs		$20,000	$20,000		$20,000		$60,000	$10,000

1.2.2.2 Web Servers

We assume a web server is a reasonably fancy desktop with lots of disk.

	2001	2002	2003	2004	2005	2006	Total	2007 (operations)
Web Server		1	0	0	2	0	3	2/3

Costs		$5,000			$10,000		$15,000	$3,300

1.2.2.3 Scheduling Tools
There are no explicit hardware costs associated with this item.

1.2.2.4 Other
There are no explicit hardware costs associated with this item.

1.2.3 Software Development Environment
This would cover licensing costs. So far, CERN has bought CMS wide licenses for all CMS computing as in, for example, objectivity. Certainly licenses for software development is important, and for software professionals the license fee is contained in their total encumbered salaries.

One can imagine licenses such as lsf licenses, etc., or licenses for any special clustering software for analysis or commodity computing. This is included in the infrastructure overhead.

1.2.4 System Administration, Monitoring and Problem Resolution Software
There are no explicit hardware costs associated with this item.

1.2.5 User Help Desk
There are no explicit hardware costs associated with this item.

1.2.6 Training Office
There are no explicit hardware costs associated with this item. Although a training office requires infrastructure, we are planning to use existing or future Fermilab facilities for this.

1.2.7 Support for Tier 2 Centers
There are no explicit hardware costs associated with this item.

1.2.8 Computer Security
There are no explicit hardware costs associated with this item.

1.3 Operations and Infrastructure

1.3.1 Software license maintenance

There are no explicit hardware costs associated with this item.

1.3.2 Software license maintenance

There are no explicit hardware costs associated with this item.

1.3.3 Infrastructure support

There are no explicit hardware costs associated with this item.

1.4 Tier 2 Regional Centers
The Tier 2 hardware configurations were chosen so as to minimize equipment and support costs and provide more flexibility than is possible at a large Tier 1 facility (though less than would be found at an average university site). Certain tradeoffs in cost, performance and mission were considered. For example, Tier 2 CPUs will be less expensive

than those used at Tier 1, but they will still cost more than the cheapest commodity processors in order to provide high performance, excellent reliability and simple maintenance.

The same considerations of performance, reliability and maintenance require other items such as high-speed LAN switches, large capacity and high-throughput data servers, uninterruptible power, dense rack mounted systems, software debugging tools, load balancing software, monitoring / management tools, and maintenance contracts. Most centers will not have Hierarchical Storage Managers (HSM) for transparent tape-disk handling because of the personnel overhead. However, one or possibly two centers might be able to take advantage of an existing HSM and its support personnel and not incur extra costs. Similarly, since the centers will be built from inexpensive CPUs, they will have more processing power than strictly needed for simulation in order to allow some of them to perform more simulations than average while other centers serve other needs.

The costing philosophy for Tier 2 hardware is that an initial purchase is made in some year, then 1/3 of that funding is used to purchase hardware for each following year, effectively replacing that hardware every three years. Other items such as maintenance and software licenses must be paid every year.

1.4.1 Prototype Tier 2 Centers

1.4.1.1 Prototype Tier 2 Center #1

The center hardware was purchased previously from FY00 funds for $350K, which is not listed here. $120K must be spent every year for replacement hardware during the prototype phase FY01–FY03.

1.4.1.1.1 Tier 2 Design

This task covers the R&D needed to design this prototype center. No funds are needed for this item.

1.4.1.1.2 Tier 2 Equipment

1.4.1.1.2.1 Hardware Procurement & Installation

This task covers the purchase and installation of the equipment.

1.4.1.1.2.2 Hardware Commissioning

It is expected that the heterogeneous hardware configuration (dual CPU boxes, data server, disks, LAN switches, power supplies, cooling and cabling) will require a commissioning phase to ensure smooth operation. Testing and reconfiguration of hardware will take place during this phase.

1.4.1.1.2.3 Hardware Support and Maintenance

This task covers all efforts related to the long-term maintenance of the equipment following the commissioning phase. Included are replacement or repair of broken equipment, testing, bringing hardware online or offline as needed, kicking broken equipment across the room, routine inspections and operating system updates.

1.4.1.1.3 Tier 2 Software Infrastructure Framework

This task covers items and activities associated with Tier 2 grid software for this prototype center, where the allowed activities are described in the introduction to this section.

1.4.1.1.3.1 Software Testing

This task covers testing of grid software developed outside the project and other software developed in support of Tier 2 operations.

1.4.1.1.3.2 Software Deployment

This task covers all items and activities related to the deployment of grid software and other software developed in support of Tier 2 operations.

1.4.1.1.3.3 Software Maintenance

This task covers all items and activities related to maintaining grid software and other software developed in support of Tier 2 operations. Possible activities include installing and testing new software releases, monitoring of performance of existing software, evaluating and comparing software, and developing tools to automate or simplify software maintenance.

1.4.1.1.4 Simulation of

1.4.1.1.5 Tier 1 / Tier 2 Integration

This task includes all items and activities related to integrating the computing facilities with that of the Tier 1 center in support of a distributed computing system.

1.4.1.1.5.1 Database replication

1.4.1.1.5.2 Object replication

1.4.1.1.5.3 Database replication

1.4.1.1.6 Tier 2 / Tier 2 Integration

This task includes all items and activities related to integrating the computing facilities with that of other prototype Tier 2 centers in support of a distributed computing system.

1.4.1.1.7 System Documentation

Documentation of Tier 2 hardware and software is needed by the physicists and staff at this center (particularly as support personnel and students are hired), by other Tier 2 sites that will be coming online later and by physicists at other CMS institutions that will be using Tier 2 services.

1.4.1.2 Prototype Tier 2 Center #2

This item covers the effort needed for the second Tier 2 Regional Center.

1.4.1.3 Prototype Tier 2 Center #3

This item covers the effort needed for the third Tier 2 Regional Center.

1.4.2 US-CMS University Tier 2 Centers

These centers will come online in FY2004, the same time as the Tier 1 center. Each Tier 2 center will have a Level 3 manager who will be responsible for ensuring that the milestones and deliverables applicable to that site are met on schedule. Centers are listed separately as they will ramp up at different times.

1.4.2.1 Tier 2 Center #1
1.4.2.2 Tier 2 Center #2
1.4.2.3 Tier 2 Center #3
1.4.2.4 Tier 2 Center #4
1.4.2.5 Tier 2 Center #5

1.5 Networking

Non-FTE networking costs fall into two general categories:

- hardware costs for network devices that support CMS systems at Fermilab;
- recurring costs for services associated with CMS network support, principally costs associated with any upgrade of Fermilab off-site network facilities in support of wide-area CMS collaboration.

The estimated non-FTE costs for networking support of CMS at Fermilab are listed below, broken down according CMS WBS item.

1.5.1 Onsite network infrastructure

A very high performance LAN switch to serve as the core CMS network switch fabric is tentatively planned for procurement in 2003. In 2001 & 2002, CMS system network support at Fermilab will be provided on existing switch fabric, although additional switch modules will need to be procured to provide switch ports for the expanding number of CMS systems on-site.

The CMS analysis, data management, and data storage systems, based in the FCC, are assumed to have gigabit network connections. Network support cost estimates are based on $500/port in 2003, $450/port in 2004, and $400/port in 2005. The per system cost estimate includes the cost of the system's gigabit switch port, plus a small amount toward support the general network infrastructure and basic network services used by CMS systems at Fermilab.

CMS desktop network connections at Fermilab are assumed to be 100Mb/s. Desktop support cost estimates are based on $250/port. The desktop cost estimate includes the cost of the switch port, a small amount toward support the general network infrastructure and basic network services used by CMS systems at Fermilab.

Year	2001	2002	2003	2004	2005	2006	Total	2007
CMS Core Switch Fabric	0	0	$30,000	$40,000	$20,000	$20,000	$110,000	$10,000
FCC networking	$20,000	$20,000	$20,000	$19,800	$29,600	$34,400	$143,800	$48,000
Desktop support	$1,250	$2,500	$6,250	$6,250	$5,000	$3,750	$25,000	$8,350
Total	$21,250	$22,500	$56,250	$66,050	$54,600	$58,150	$278,800	$66,270

1.5.2 Off-Site Networking

Network support for off-site CMS collaboration is generally assumed to be provided by shared wide-area HEP network facilities, including the LHC trans-Atlantic link, ESnet, and the Internet-2 networks. The Fermilab ESnet link, currently OC3, is assumed to be OC12 by 2002, as part of general upgrade of Fermilab wide-area network capabilities. Additional bandwidth capacity to meet CMS Tier 0/Tier 1 and Tier 1/Tier 2 bandwidth needs is projected as a recurring cost, starting in 2004. A one-time hardware cost for network equipment to support the additional bandwidth is projected for 2004. Modest hardware costs are projected for procurement of servers to support network-level functions of data grid services, starting in 2004.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Off-site access hardware	0	0	0	30000	0	0	30000	10000
Leased line costs	0	0	0	250000	250000	250000		250000
Grid network support servers	0	0	0	10000	5000	5000	20000	5000
Total	0	0	0	290000	255000	255000	50000	265000

1.5.3 Network Security

Hardware costs for CMS LAN firewall protection, intrusion detection, and related network-level security devices are projected in 2003, when the core CMS switch is procured. Upgrade and enhancement of those devices is projected through 2006. Modest hardware costs are projected for procurement of servers to support security functions of data grid services, starting in 2004.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Network security Hardware	0	0	0	20000	10000	10000	40000	10000
Grid security support servers	0	0	0	10000	5000	5000	20000	5000
Total	0	0	0	30000	15000	15000	60000	15000

1.5.4 Network Performance Monitoring

A wire-rate gigabit LAN analyzer and associated packet capture probes would be procured in 2003, with modest upgrades or enhancements in the following years. Modest hardware costs are projected for procurement of servers to support wide-area performance monitoring functions of data grid services, starting in 2003.

Year	2001	2002	2003	2004	2005	2006	Total	2007
Monitoring tools & hardware	0	0	0	40000	10000	10000	60000	20000
Grid WAN perf. Monitor servers	0	0	0	10000	5000	5000	20000	5000
Total	0	0	0	50000	15000	15000	80000	25000

1.6 Computing and Software R&D

1.6.1 Distributed Data and Computing Test Beds

This covers a test bed to study distributed data access and data access monitoring. Results from this will be used to validate the MONARC simulation tool. The testbed will support distributed data handling software and monitoring as well as studying data access patterns.

This does not require a lot of computing. A large size server with good network connectivity (e.g. Gb Ethernet) is fine. This machine must be for the exclusive use of these tests in order to monitor everything that goes on with them, however. We have purchased a DELL 4cpu linux box in 2000 we will use for this, but we believe we may need a real server class machine, which we plan to purchase in 2001. The dell machine cost $25k with ¼ TB h/w raid. We are assuming in 2001 a server class machine with order TB of disk will cost about $30k.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Boxes	1	1						
Cost	$30,000	50,000	0	0	0	0	$80,000	0

1.6.1.1 Hardware Testbed system design
There are no explicit hardware costs associated with this item.
1.6.1.2 Hardware testbed system
There are no explicit hardware costs associated with this item.
1.6.1.3 Distributed data management
There are no explicit hardware costs associated with this item.
1.6.1.4 Distributed Data Access
There are no explicit hardware costs associated with this item.
1.6.1.5 Monitoriing Tools
There are no explicit hardware costs associated with this item.

1.6.2 R&D in Computing Hardware for T1/T2 Computing
The cost in this item comes from computing costs for investigating different kinds of computing technology. This is assumed to be bought during 2000-2002 as new things come in.

1.6.2.1 Technology Tracking and Testing
There are no hardware costs associated with this item.

1.6.2.2 Technology Investigation and Deployment

1.6.2.2.1 Data Intensive Computing

Data intensive computing refers to computing used for analysis needed a very high I/O. These machines must have excellent inter-connectivity and lots of I/O capability. Much of this can probably be achieved with a linux cluster, however we may need a more expensive SMP –type solution. The costs here are for a linux cluster + integration costs.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Linux Boxes	10	20	20	0	0	0	50	0
Cost	$50,000	$100,000	$100,000	0	0	0	$250,000	0

1.6.2.2.2 CPU Intensive Computing

We believe we know how to do this, and this makes up a relatively small portion of the total CPU for the Regional Center. Basically we need to keep evaluating new boxes as they come along, but we don't need to build up a large capability here. These requirements could change depending on new data challenge milestones.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Cost	0	50,000	50,000	0	0	0	$100,000	0

1.6.2.2.3 Data Access and Distribution Servers

Here we don't know if linux machines will work or if we need higher end servers. Experience from CMS MC production suggests one server class machine for each 20 cpus in order to get the I/O for the MC production. We need lots of disk as well. The cost is for larger SMP servers, but we have bought already some smaller linux servers in order to compare.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Data server boxes	1	2	2	0	0	0	5	0
Disk (TB)	6	12	6	0	0	0	24	0
Hardware Cost	$220,331	$307,636	$163,277	0	0	0	$691,245	0

1.6.3 Prototype Fully Functional Regional Center

The fully functional regional center should be 5% of the final tier 1 center. It should come online end of FY 2003. Listed here is an approximate cost for the additional incremental costs to finish it in addition to the R&D systems that can be used as parts.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Total	0	0	$100,000	0	0	0	$100,000	0

1.6.4 Simulation Software

There are no explicit hardware costs associated with this item.

1.6.5 Simulation Software

There are no explicit hardware costs associated with this item.

1.6.6 Simulation Software
There are no explicit hardware costs associated with this item.

1.6.7 Data Import/Export R&D

These are incremental costs for R&D for tape input/output. We assume we can use the same AML/J we are using to support the construction project.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
# tapes	0	60	0	0	0	0	60	0
Tape costs	0	$4,200	0	0	0	0	$4,200	0
Drive cost	0	$25,000	0	0	0	0	$25,000	0
Total	0	$29,20	0	0	0	0	$29,200	0

1.7 Detector Construction Phase Computing

1.7.1 Computing

We assume we can support the detector construction project computing on existing central FNAL systems as well as CMS systems we already have, and the R&D systems as we replace them. We add no new costs here.

1.7.2 Data Storage
This item covers data storage to support physicists doing detector or physics studies during the construction project.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Tape costs	0	17500	15000	0	0	0	$32500	0
Drive cost	$50,000	0	0	0	0	0	$50,000	0
Maintenance	$3,000	$4,500	$6,000	$6,000	0	0	$19,500	0
Total	$53,000	$22,000	$21,000	$6,000	0	0	$102,000	0

1.7.3 Network based file interchange
There are no hardware costs associated with this item.

1.7.4 Tape import/export system
This covers the cost of refitting the AML/J, exporting 50 tapes in 2002 and 100 tapes in 2003 and the maintenance on the library and tape drives.

Year	2001	2002	2003	2004	2005	2006	Total	2007(ops)
Total	$82,500	$12,500	$27,500	0	0	0	$122,500	0

1.8 Support for FNAL based computing
1.8.1 Desktop systems

1.8.1.1 Installation
There are no explicit hardware costs associated with this item.

1.8.1.1.1 System specification, design and installation

There are no explicit hardware costs associated with this item. FNAL costs are in the total encumbered salaries of computing professionals, or come from the base program support of scientists. Outside users who come to FNAL are expected to bring their own desktops (or laptops), which must conform to the FNAL standard.

1.8.1.1.2 Host server systems

We are assuming we support 100 desktops, 50 FNAL, 50 outside users by the year 2005. This ramps up from 10 in 2000 to 100 in 2005 and onwards. We are not buying the actual desktops but we are buying disk servers to support them. Run 2 experience suggests 1 server class machine for every 20 desktops.

	2001`	2002	2003	2004	2005	2006	Total	2007 (operations)
#desiktops (integrated)	10	15	20	30	60	100	100	100
# new servers		1		1	1	2	4(+1)	1
costs	$10,000	$25,000	$15,000	0	$25,000	0	$75,000	$12,500

1.8.1.1.3 Backup systems

This item covers backups for user areas, etc. I am assuming a reasonable system is about $10k/50 machines.

	2001`	2002	2003	2004	2005	2006	Total	2007 (operations)
#desiktops (integrated)	10	15	20	30	60	100	100	100
# new backup systems	0	1	0	0	0	0	1	0.25
costs	0	$37,000	0	0	0	0	$37,000	$9,250

1.8.1.2 Support

There are no explicit hardware costs associated with this item.

1.8.1.3 User data storage and archiving

This is a backup and archiving system for user data from their desktops. Because the server class machines will hold a lot of data disk, we should not back it up automatically but allow users to backup and archive what they want. The minimum would be a small stacker on each server class machine, which should cost on the order of $10k/machine.

	2001	2002	2003	2004	2005	2006	Total	2007 (operations)
costs	0	0	0	$85,000	0	0	$85,000	$10,000

1.8.2 Remote control room

	2001	2002	2003	2004	2005	2006	Total	2007 (operations)
costs	0	0	$20.000	$10,000	$10,000	$10,000	$50,000	$10,000

TOTAL HARDWARE COST FOR THE USER FACILITY PROJECT

Here we summarize the total cost for the User Facility Project. Each of the WBS items (except for 1.4) also has an additional overhead of 16% which is charged by Fermilab.

WBS #	2001	2002	2003	2004	2005	2006	Total
1.1 T1 Regional Center	0	0	0	2,648,908	2,838,675	2,595,398	8,082,981
1.2 System Support	0	29,000	23,200	0	34,800	0	87,000
1.3 O&M	0	0	0	0	0	0	0
1.4 T2 Regional Centers	280,000	470,000	590,000	1,120,000	1,750,000	1,750,000	5,960,000
1.5 T1 Networking	24,650	26,100	65,250	543,924	449,152	458,838	1,567,914
1.6 Computing R&D	348,384	622,730	479,402	0	0	0	1,450,516
1.7 Current Support	157,180	40,020	56,260	6,960	0	0	260,420
1.8 Local Comp. Supp.	11,600	71,920	139,200	23,200	52,200	23,200	321,320
Total	**821,814**	**1,259,770**	**1,353,312**	**4,342,992**	**5,124,827**	**4,827,436**	**17,730,151**

BIBLIOGRAPHY

[1] *Regional Centres for LHC Computing*, Report of the MONARC Architecture Group